Contemporary Life Science Research (CLSR)
ISSN (Online): **XXXX-XXX**
Volume 01 | Issue 02 | October 2026
DOI:

Integrative Functional Genomic Analysis of CRISPR Screening Data to Prioritize Host Dependency Factors in Coronavirus Infection

Dr Anchit Raja[1] , Dr. Arti Pandoh Gupta[2]

[1]2nd Year PG Radiology, Radiology, GMC Jammu, University of Jammu, Email Id: Anchitraja.rap@gmail.com, Orcid Id: 0009-0008-6381-0920

[2]Principal, physics, Government Degree College Batote, District Ramban, Jammu and Kashmir, University of Jammu J& K 180011, Email ID: Arti181423@gmail.com, ORCID ID: 0009-0009-4634-4785

Received Date: **12-06-2026** | Revised Date: **19-08-2026** | Accepted Date: **21-09-2026** | Published Date: **13-10-2026**

ABSTRACT
Electrophoretic Coronavirus infection depends on multiple host cellular factors that regulate viral entry, replication, and intracellular processes. Identifying reliable host dependency factors is essential for understanding coronavirus–host interactions and supporting future antiviral research. The present study investigated potential host dependency factors associated with SARS-CoV-2 infection through comparative analysis of publicly available genome-wide CRISPR-Cas9 screening data obtained from the NCBI Gene Expression Omnibus (GSE178699). The analysis included 18,053 non-control gene labels across five cellular models: Vero, Huh7, Calu-3, Caco-2, and HEK293A_T. Gene-level log2 screening differences, cross-model consistency, guide coverage, and exploratory prioritization scores were evaluated using descriptive statistical approaches and Pearson correlation analysis. ACE2 demonstrated the highest exploratory priority score (3.661), with positive screening differences exceeding the selected threshold in three cellular models. Additional candidates, including GYPB, KRTAP1-1, HNF1B, and PCBD1, exhibited variable screening responses. The strongest correlation was observed between Vero and Calu-3 cells ($r = 0.247$), indicating limited similarity across cellular models. These findings highlight substantial heterogeneity in coronavirus-associated screening profiles and emphasize the importance of cross-model comparisons for candidate prioritization. Provisional experimental contrasts, the absence of formal statistical significance testing, and limited functional validation restrict definitive biological interpretation. Overall, the study provides a preliminary framework for prioritizing potential coronavirus host dependency factors and supports further functional genomic investigations and experimental validation.

Keywords: SARS-CoV-2, CRISPR-Cas9 screening, Host dependency factors, Functional genomics, Gene prioritization.

1. Introduction

Coronavirus infections are an important health threat due to the potential for respiratory illness and easily transmissible among populations. The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) requires host cell machinery for viral entry, replication and propagation. It is therefore crucial to gain understanding of the biology of the virus to identify cellular processes that facilitate virus infection and potential therapeutic targets (V'kovski et al., 2021). The entry of viruses is mediated mainly by binding of spike protein to angiotensin-converting enzyme 2 (ACE2); spike protein activation is mediated by transmembrane serine protease 2 (TMPRSS2) (Hoffmann et al., 2020). Once the virus enters the cell, it takes over its cellular resources to help with RNA synthesis and viral protein production, resulting in complex transcription products that facilitate viral replication (Kim et al., 2020). In understanding these processes, it will also be important to identify cellular host factors that can be targeted for the development of host-directed antiviral approaches that have broader therapeutic potential (Baggen, Vanstreels, et al., 2021).

Clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 screening has become an important tool in functional genomics to identify genes that are associated with viral infection. This is a technology that allows to systematically disrupt genes and study the resulting alterations in the susceptibility of the cells to infection. CRISPR-based strategies have played a significant role in elucidating coronavirus–host interactions and identifying cellular factors which can affect viral replication (Cui et al., 2023). Cellular trafficking and endosomal pathways, as well as other host-dependent processes, have been identified as critical for SARS-CoV-2 infection through genome-wide knockout screening (Daniloski et al., 2021). Further screening investigations have shown that host genetic factors affect the outcomes of infection by multiple mechanisms in their biological actions (Wei et al., 2021). Genome-wide screening offers a wider genetic coverage than traditional candidate gene strategies and can identify novel host factors. But genetic associations have to be interpreted with caution given the differences in screening procedures and experimental conditions (See et al., 2024).

Even with these advances, it is still difficult to identify reliable host dependency factors, due to the fact that screening results often differ between cellular models and experimental conditions. CRISPR screening of the entire genome has revealed the host factor transmembrane protein 106B (TMEM106B) to be involved in SARS-CoV-2 infection, beyond known viral entry pathways (Baggen, Persoons, et al., 2021). Other studies have determined other host factors that play a role in coronavirus infection, but their specific functions may vary depending on the experimental system (Grodzki et al., 2022). Virus–host interactions are complex, as shown by the identification of common host factor networks and a coronavirus-specific network by comparative screening studies (Schneider et al., 2021). Through genetic screening of SARS-CoV-2 and common cold coronaviruses, other cellular dependencies have also been identified as virus-specific, and others as conserved (Wang et al., 2021). These outcomes propose the importance of systematic comparisons to identify the reproducible genetic signals and to differentiate them from model-dependent observations.

To overcome inconsistencies, an integrative functional genomic analysis could be used that integrates screening and biological data. Previous cross-model investigations have revealed ACE2 as a key shared dependency and shown that there are large differences across the other candidate host factors (Chan et al., 2021). Protein interaction mapping has also identified numerous interactions between SARS-CoV-2 proteins and host cell components, further highlighting the importance of mapping the interaction of proteins in biological networks as a means of biological interpretation (Gordon et al., 2020). Comprehensive multi-omics studies reveal antiviral host factors and pathways, highlighting the promise of integrating complementary genomic information (Hou et al., 2024). Genome-wide screening targeting viral entry has shown that different host genes affect the susceptibility of a cell to the infection by SARS-CoV-2 (Zhu et al., 2021). Screening-effect size, consistency across models and functional significance should be taken into account for reproducible prioritization of candidate genes.

Hence, the goal is to identify and compare candidate host dependency factors based on the publicly available SARS-CoV-2 CRISPR screening data, investigate their consistency in cellular models, and prioritize potential host factors based on the integrative analysis of the host screening data. The study also investigates the importance of functional pathways as well as protein interactions in prioritized candidates, which is crucial for further biological validation and antiviral studies in the host.

2. Methodology

2.1. Study Design and Data Source

In this study the secondary data based computational research design was used to explore the possible host dependent factors of the coronavirus infection from genome wide CRISPR-Cas9 screening data. Publicly

available genome-wide CRISPR screening data were retrieved from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO), accession number GSE178699 (Brown et al., 2022). Systematic Genome-Scale Identification of Host Factors for SARS-CoV-2 Infection Across Models Yields a Core Single Gene Dependency. The first screen logs contained 71,090 guide level observations over 35 samples and were converted to a gene level analytical log consisting of 18,056 gene labels. Comparative screening responses, consistency across different types of cell models and preliminary prioritization of cell dependency factors were investigated. The research process depended entirely on secondary data which was publicly available, no new experimental procedures and participant recruitment were carried out.

2.2. Data Selection and Eligibility Criteria

The data analyzed were chosen based on their relevance in the context of SARS-CoV-2 infection, the presence of CRISPR screening measurements, and identification of targeted genes. The gene identifiers, read counts from the guides and information about the experimental cellular models in the records were deemed appropriate for analysis. There were 18,056 labels in the gene level file prepared, including 18,053 gene labels (excluding control) and 3 control labels. For documentation, but not prioritization of candidates, the control entries LacZ, luciferase and EGFP were also kept. All five of the available cellular models were used for data to assess the variation in screening responses. The information available from GSE178699 was the only data analysed and no independent screening studies were included. Therefore, comparisons did not reflect variation between distinct investigations, but rather within specific investigations that were the same cell models.

2.3. Data Preprocessing and Normalization

The original guide-level read counts were cleaned up for an analytical file at the gene-level. The total sequencing read count, the number of guides with zero count and the percentage of the guides with zero count were used to summarize the sample level quality characteristics. Guide counts were scaled based on the total number of counts per sample, and presented as counts per one million counts. Because of zero count observations a pseudocount of 0.5 was added prior to normalization and logging transformation. Log2 abundance differences were determined by normalization of log2 abundance of the designated reference samples to the designated comparison samples. Guide level differences were then summarized at the gene level, with medians of the guide-level differences of the guides that targeted the same gene. An analytical file that included five cell-model-specific screening measurements, five gene-level summary scores, and markers of cross-model consistency, was created. Comparisons with the labels were treated as provisional since they were not independently verified against the original sample metadata.

2.4. Identification of Candidate Host Dependency Factors

The gene-level screening measurements for five cellular models were used to investigate candidate host dependency factors. The screening effect amplitude, direction and uniformity across the cellular models were analyzed. Positive and negative log2 screening differences were analyzed separately to see if there were differences in screening responses to the genes. Genes that exhibited comparatively strong positive screening signals were considered to have a difference greater than positive log2 difference of 1, and those with positive screening signals less than negative log2 difference of 1 were considered positive screening signals that were not strong. The thresholds were not used for statistical hypothesis testing, but for exploratory classification. The number of CRISPR guides that were used to measure each gene level was also noted to aid interpretation of the results. Genes with less supporting evidence in the form of guides were deemed to be less well supported. The list of the identified candidates was treated as a list of genes that were considered to be preliminary screening associated genes, since the available analytical file did not include formal gene-level significance estimates and experimentally validated dependency classifications.

2.5. Cross-Screen Consistency Analysis

The consistency of the cross-screen was evaluated by comparing the screening measurements at the gene level between the Vero, Huh7, Calu-3, Caco2 and HEK293A_T cellular models. For each gene, the number of models that showed positive screening differences, over the descriptive threshold, was noted. The fraction of the five cellular models that met this criterion (cross screen consistency) was reported as a percentage. Median and mean log2 differences in screening were also analyzed to provide a general overall magnitude and direction of responses to screening. Genes with similar positive results in multiple models were identified from those with model-specific and/or conflicting results. Screening counts were kept separately for positive and negative, to detect differences in the direction of the effects. This comparison was designed to provide a description of the reproducibility in experimental cellular systems, without imposing the same biological properties or infection conditions on all models.

2.6. Functional Genomic Interpretation

All candidate genes screened in this study were clustered according to their functional interpretation in

relation to the screening response and biological relevance to coronavirus infection. Biological processes, molecular functions and cellular pathways were further examined, with candidates identified for further screening through gene-level screening profiles. Special emphasis was placed on identifying genes that have been consistently screened, and those that have been found to respond to screening in highly variable ways in different cellular models. The functional enrichment performed by the use of biological terms from the Gene Ontology (GO) or pathway databases already built was regarded as an extra part of the suggested integrative approach, alongside with protein–protein interaction analysis. The analysed file uploaded however did not include results of functional enrichment nor protein interaction. No attempt has been made to consider pathway enrichment, molecular interaction networks and identification of biological modules as concluded analyses.

2.7. Gene Prioritization and Statistical Analysis

The variables used for exploratory ranking were in the prepared analytical file and were used to prioritize genes. To account for the significance of screening differences in the cellular models, and the absolute size of the median gene-level screening response, the ranking approach was used. Each gene was assigned an exploratory priority score, which was then used to prioritize each of the candidate genes based on the relative screening evidence. Descriptive statistical measures were frequencies, percentages, mean screening differences, median screening differences and cross-screen consistency fractions. The screening results were summarised by the five cellular models and the distributions at the level of genes were found to be appropriate for descriptive comparisons and graphical presentation. The results of genes for which screening and guide coverage differed were treated with some caution. The file did not include any formal p-values, confidence intervals or false-discovery-rate adjusted significance estimates.

2.8. Ethical Considerations

In the present investigation, only publicly available secondary genomic screen data, from the NCBI Gene Expression Omnibus, were used. No new laboratory experiments, no recruiting of human participants and no personally identifiable information was collected. The original data source was recognized to ensure transparency and integrity of research. Since the analysis was based on re-use of publicly available screening records, there were no further participant related ethical risks.

3. Results

3.1. Characteristics and Distribution of CRISPR Screening Data

Data analyzed comprised 18,056 gene labels, with three being control labels and the remaining 18,053 being non-control labels. For five cellular models (Vero, Huh7, Calu-3, Caco-2 and HEK293A_T), gene-level exploratory screening scores were available. Considerable variation was observed in gene-level screening differences across the five cellular models. The lowest mean log2 screening difference was observed in Vero cells (−1.363), followed by Calu-3 cells (−1.150), whereas Caco-2 cells exhibited a positive mean difference (0.158). Vero and Calu-3 cells showed substantially higher numbers of genes with negative screening differences below −1, comprising 11,059 and 9,966 genes, respectively. The number of the differences in exploratory screening profiles were the largest in Caco-2 with a majority of the differences being positive differences with a value greater than 1, showing significant variation between the two cellular models. The results are briefly summarized in Table 1 and shown graphically in Figure 1.

Table 1. Distribution of Gene-Level CRISPR Screening Differences Across Cellular Models

Cell Model	Genes (n)	Mean log2FC	Median log2FC	Positive Genes (n)	Negative Genes (n)
Vero	18,053	-1.363	-1.210	65	11,059
Huh7	18,053	-0.050	-0.034	1	59
Calu-3	18,053	-1.150	-1.094	30	9,966
Caco-2	18,053	0.158	0.157	379	102
HEK293A_T	18,053	-0.071	-0.064	5	32

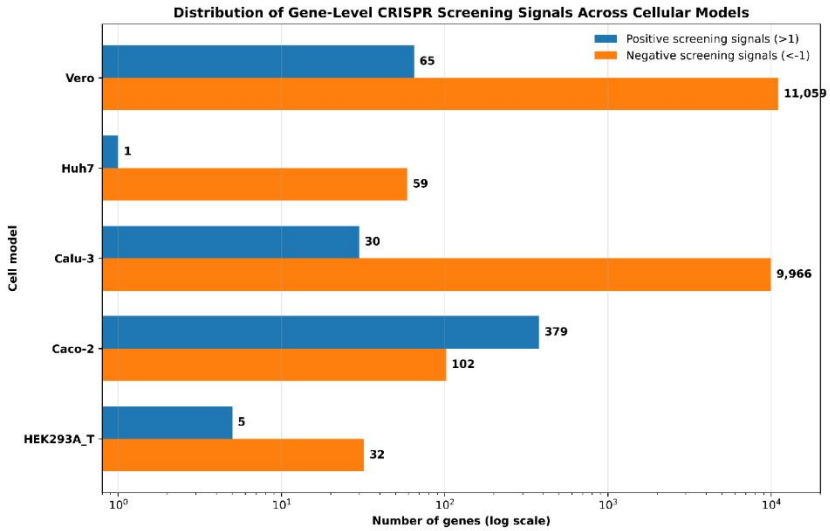

Figure 1. Model-Specific CRISPR Screening Signal Distribution

3.2. Identification and Distribution of Candidate Host Dependency Factors

A relatively small number of genes had positive log2 differences greater than 1, as shown by the screening analysis. Of the 18,053 non-control genes, the number exceeding this threshold was 17,577 (97.36%) in none of the cellular models, and 473 (2.62%) in only one cellular model. Two genes were found to exceed positive thresholds in two models and one gene exceeded positive thresholds in three models. 14,568 genes (80.70%) had negative screening differences in at least one model. In addition, 391 genes (2.17%) showed both positive and negative differences with the threshold exceeding across different cellular models. As indicated in Table 2 and Figure 2, positive screening responses were infrequent and the individual response to the gene differed greatly among the experimental models.

Table 2. Cross-Model Distribution of Candidate Screening Signals

Screening category	Number of genes	Percentage (%)
No model with positive difference >1	17,577	97.36
Exactly one model with positive difference >1	473	2.62
Exactly two models with positive difference >1	2	0.01
Exactly three models with positive difference >1	1	0.01

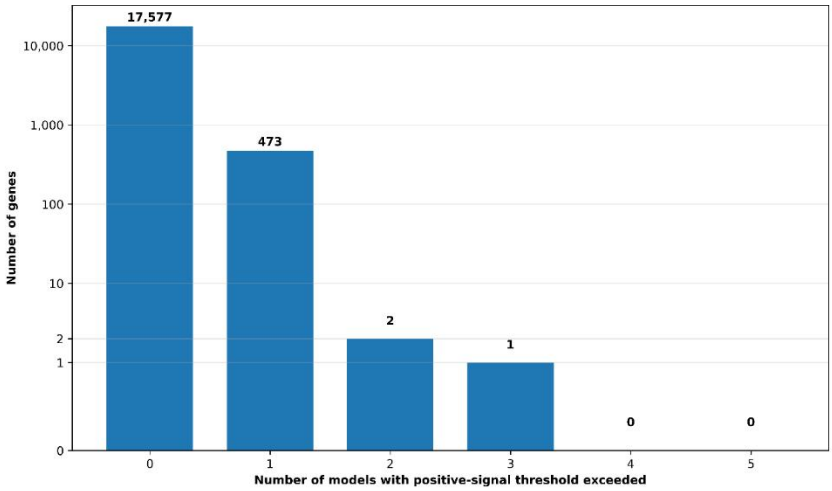

Figure 2. Cross-Model Consistency of Positive Screening Signals

3.3. Cross-Screen Consistency and Reproducibility of Candidate Host Factors

The cross-screen consistency analysis demonstrated significant inter-model variability in the response of the individual genes. ACE2 had the highest exploratory priority score (3.661) with positive screening differences >1 in 3 models and a median log2 difference of 2.645. GYPB was second, followed by KRTAP1-1, which was only represented by one CRISPR guide and had screening directions in opposite

directions in models. Two more genes, HNF1B and PCBD1, had four guides each, but neither of these hit the positive threshold in all four models. Other genes such as MED12, EP300 and SMAD4 showed smaller medians of screening differences. The results suggest that exploratory prioritization was related to the magnitude of screening responses and their spread over the different models of cells. The top candidates are listed in the Table 3 and their profiles modeled in the Figure 3.

Table 3. Highest-Ranked Candidate Genes Based on Exploratory CRISPR Screening Scores

Rank	Gene	Guides	Median log2 difference	Models >1	Priority score
1	ACE2	4	2.645	3	3.661
2	GYPB	1	0.929	2	2.232
3	KRTAP1-1	1	0.005	2	2.001
4	HNF1B	4	0.630	1	1.157
5	PCBD1	4	0.543	1	1.136
6	DEFB130	3	0.483	1	1.121
7	HIST1H3G	1	0.461	1	1.115
8	POMZP3	1	0.457	1	1.114
9	MED12	4	0.411	1	1.103
10	EP300	4	0.328	1	1.082

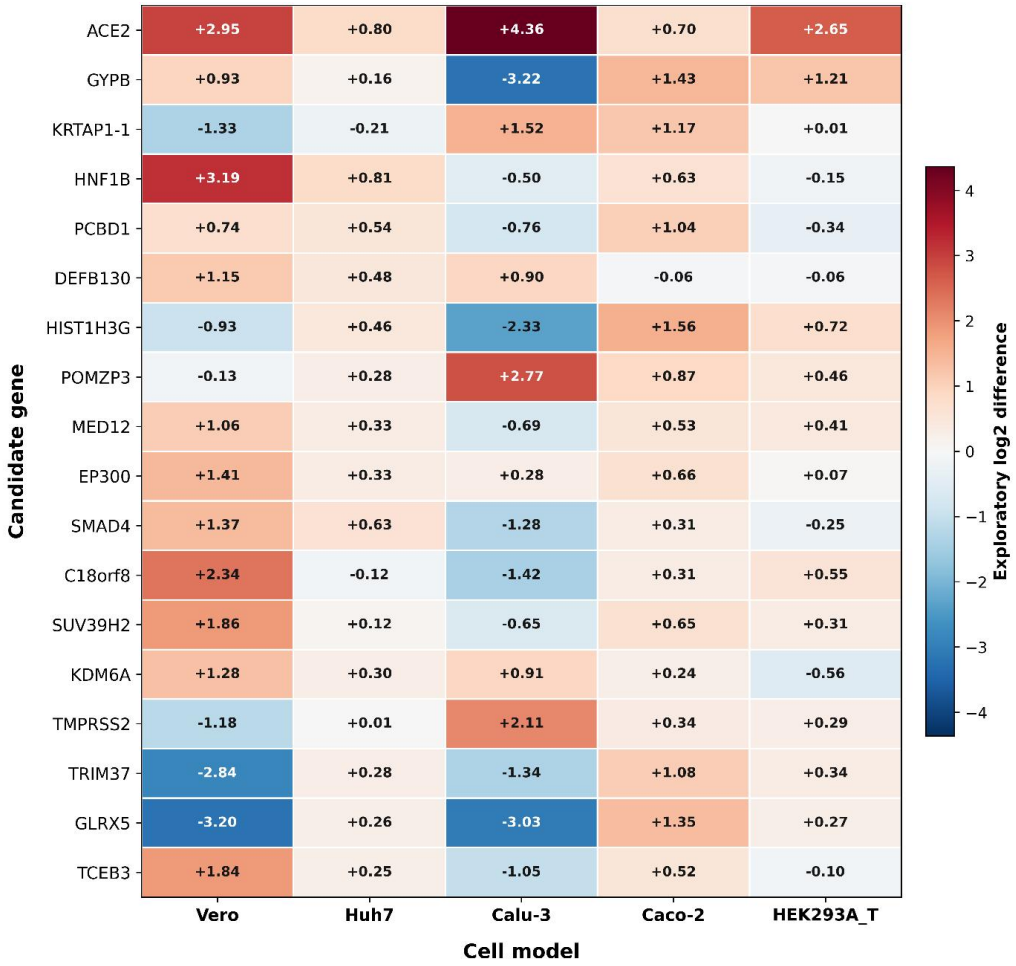

Figure 3. Heatmap of Screening Differences Among Prioritized Candidate Genes

3.4. Correlation Analysis of Gene-Level Screening Responses Across Cellular Models

This correlate gene-level screening responses among cellular models. The consistency of the differences between the five cellular models was checked using the Pearson correlation analysis for the gene level screening. The highest correlation was seen between Vero and Calu-3 cells (r = 0.247), and Vero and Huh7 cells (r = 0.192). A lower positive correlation value was found between Huh7 and Calu-3 cells (r = 0.132). Correlation with Caco-2 was near zero (between 0.004 and 0.014) and no significant correlations could be

established between HEK293A_T and the other models. The results suggest only a partial similarity in the overall profile of the exploratory screening profiles for the different cellular systems examined. The correlation coefficients are found in Table 4 and graphically displayed in Figure 4.

Table 4. Pearson Correlation Matrix of Gene-Level CRISPR Screening Differences

Cell model	Vero	Huh7	Calu-3	Caco-2	HEK293A_T
Vero	1.000	0.192	0.247	0.014	−0.038
Huh7	0.192	1.000	0.132	0.014	−0.017
Calu-3	0.247	0.132	1.000	0.004	−0.024
Caco-2	0.014	0.014	0.004	1.000	0.008
HEK293A_T	−0.038	−0.017	−0.024	0.008	1.000

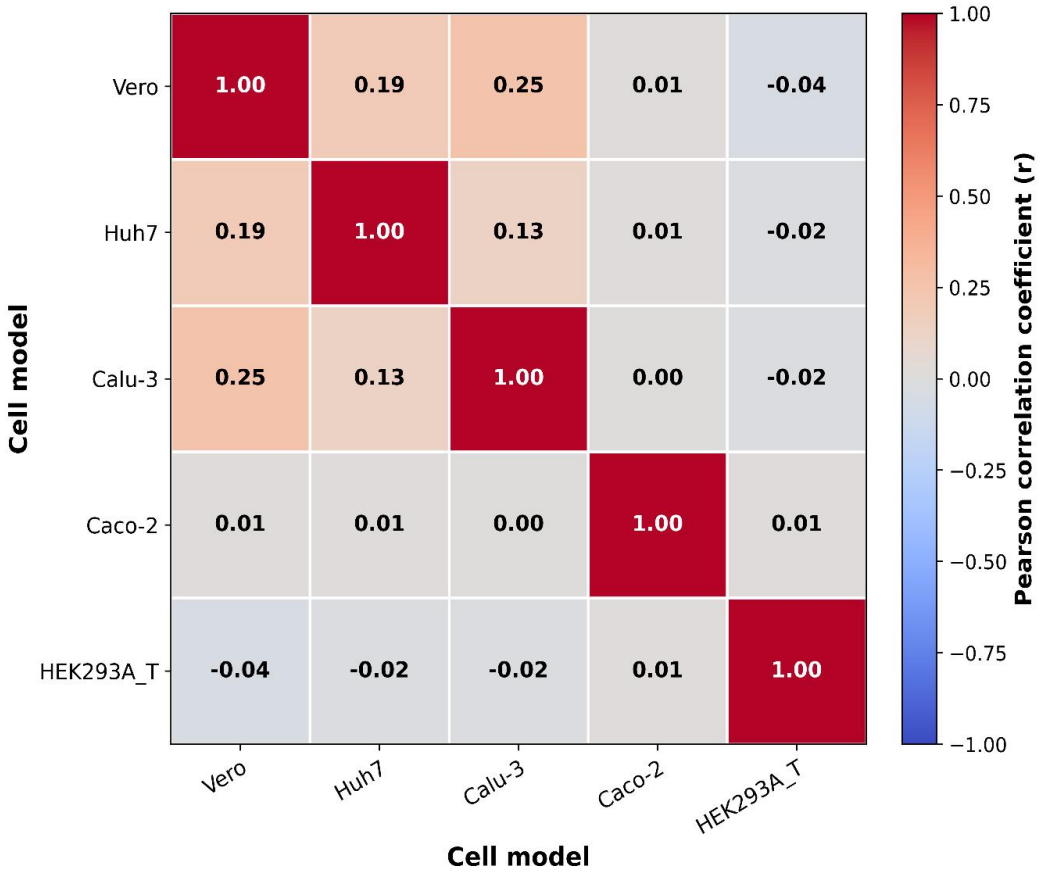

Figure 4. Correlation Heatmap of CRISPR Screening Responses Across Cellular Models

3.5. Integrated Exploratory Gene Prioritization and Guide Coverage

The integrated exploratory ranking was shown to give the highest priority score for the gene ACE2, followed by GYPB, KRTAP1-1, HNF1B and PCBD1. The distribution of the prioritization scores exhibited a large gap between ACE2 and other competitors. Further differences in the quality of the evidence available for screening were found in the guide coverage analysis. Several highly-ranked genes had only one guide, whereas ACE2, HNF1B, PCBD1, MED12, and EP300 had four guides. Of all the non-control gene set, 231 genes were targeted by less than two guides. The magnitude of scores on the screening tools and the coverage are shown to have to be taken together when interpreting candidate rankings. Table 5 shows the screening difference of selected candidates on each cell.

Table 5. Cell-Model-Specific Screening Differences of Selected Candidate Genes

Gene	Guides	Vero	Huh7	Calu-3	Caco-2	HEK293A_T
ACE2	4	2.952	0.804	4.362	0.698	2.645
GYPB	1	0.929	0.158	−3.221	1.432	1.214
KRTAP1-1	1	−1.330	−0.213	1.519	1.174	0.005
HNF1B	4	3.186	0.812	−0.504	0.630	−0.152

PCBD1	4	0.742	0.543	−0.761	1.044	−0.341
DEFB130	3	1.149	0.483	0.901	−0.055	−0.064

4. Discussion

In the present study the genome-wide CRISPR screening data were analysed to find and prioritize possible host dependency factors associated with SARS-CoV-2 infection. ACE2 had the highest exploratory priority score (3.661) and positive screening differences greater than the threshold in three cellular models. This result is in line with the known role of ACE2 in coronavirus entry, since structural studies have shown that the spike protein receptor-binding domain (RBD) of coronaviruses specifically binds to ACE2 (Lan et al., 2020). Further structural information has established SARS-CoV-2 as having a biologically relevant molecular recognition of the human ACE2, which is a major entry receptor (Yan et al., 2020). Other top-ranked candidates were GYPB, KRTAP1-1, HNF1B, and PCBD1, but these had different screening responses in the various cellular models. Host dependence has been shown to be indirect regulation of the expression of viral entry mechanism, as previous investigation showed that GATA6 regulated the expression of ACE2 (Israeli et al., 2022). Also, it is known that in certain experimental environments TMEM106B is a receptor for entry of viral pathogens that do not rely on ACE2 (Baggen et al., 2023), suggesting that the mechanisms by which coronaviruses enter cells may vary between the different cellular systems.

The present analysis revealed a wide range of screening responses in the various cellular models. The correlations between the cells were generally highest between Vero and Calu-3 ($r = 0.247$) and were near zero for the other cells. The results suggest low similarity in gene level screening to explore. The effect of cell properties on screening results was revealed in previous bidirectional CRISPR studies, which found that mucins regulate SARS-CoV-2 infection (Biering et al., 2022). The interaction of coronavirus host proteins has also been studied and common as well as virus-specific host factors have been identified (Hoffmann et al., 2021). Comparative genome-wide screening studies between different coronaviruses have shown that the viral host dependency is dependent on the virulence of the virus and experimental conditions (Rebendenne et al., 2022). Furthermore, candidate entry receptors were identified via CRISPR activation screening, which might not be identified by conventional knockout screening (Zhu et al., 2022). Hence, the variation observed could be due to differences in experimental models, screening strategies and perturbation of genes.

The priority of host factors identified from the screening patterns indicate that it will be necessary to consider the function of the cells in addition to the effects of individual genes. Intracellular transport, membrane remodeling, RNA processing, and processes associated with replication are all processes of infection by coronavirus. Intracellular trafficking mechanisms were found to be crucial for β-coronaviruses viral egress in the previous studies (Ghosh et al., 2020). To date, the associations of SARS-CoV-2 proteins with a number of cellular components of the host cell have been identified using proximity mapping, which highlights the importance of molecular interaction information for biological interpretation (May et al., 2022). The effects of RNA–protein interactions have also been extensively investigated for the interaction of viral RNA with host proteins involved in regulation of the host cell (Schmidt et al., 2021). Structures associated with the membrane are also required for replication of the viral RNA and organization of the virus (Snijder et al., 2020). The enriched biological pathways and protein interaction networks, however, were not independently determined in the present analysis. As a result, these mechanisms explain the prioritized genes in biological terms, and are not validated by experiments.

A list of prioritized host-associated genes could be helpful in selection of target genes for future studies in host-directed antiviral research. A good screening response alone is not sufficient to prove suitability for treatment. In previous studies on CRISPR, DAXX was found to be a restriction factor in SARS-CoV-2 infection, highlighting the possibility of host proteins having either negative and/or positive effects on infection (Mac Kain et al., 2022). Likewise, proteins with intracellular localization, such as PLAC8, have been associated with SARS-CoV-2 entry, and thus, the susceptibility to infection may be influenced by these proteins (Ugalde et al., 2022). Structural studies of coronavirus replication organelles have revealed structural properties of the membrane associated molecular features that could be leveraged for future studies on developing antiviral drugs (Wolff et al., 2020). In addition, the levels of ACE2 expression are different in different cell types and can also be affected by interferon mediated responses, so that host-directed interventions need to be carefully evaluated biologically (Ziegler et al., 2020).

An important advantage of this study is that labels for 18,053 non-control genes were compared across five cellular models in an identical manner, with labels prioritized by exploration. However, the study was limited to one screening study that was publicly available and the experimental comparison assignments

need to be confirmed independently. Some of the most highly-ranked genes were only screened with a single guide, limiting the confidence of their screening evidence. Formal statistical significance testing and functional enrichment analysis, protein interaction analysis, and independent experimental validation were not performed. These constraints are used for interpreting the results to the preliminary prioritization of candidates.

Experimental contrasts should be confirmed by future studies and candidate genes should be tested using independent experimental systems. Functional enrichment and protein interaction analyses would help in the biological interpretation of prioritized candidates. Further research with physiologically relevant respiratory cell models, various strains of coronaviruses and specific gene perturbation experiments is required to establish the identified screening associations are reproducible, host dependencies.

5. Conclusion

In the present study showed factors of potential host dependency associated with SARS-CoV-2 infection by analysing genome-wide CRISPR screening data publicly available. The results showed that the level of screening responses at the gene level was quite different in five different cell models, thus underlining the complexity of the interactions between coronaviruses and their hosts. Of the candidates that were evaluated, ACE2 had the highest score on exploratory prioritization and had positive screening signals in several cellular models, and is known to be relevant in viral entry. Other candidate genes screened showed inconsistent results, and may be associated with the model type and need further investigation. The relationship between the different screening profiles were generally poor, as found by cross-model correlation analysis, which suggests that not many of the screening profiles are consistent across the various cellular systems. The results highlight the need to take into account the effect size of the screenings, their direction and their reproducibility across models when identifying potential host dependency factors. The study also shows how well secondary CRISPR screening results can be used for comparative functional genomic studies and initial candidate identification. The lack of statistical significance testing, functional enrichment analysis and experimental validation, however, hinders the biological interpretation of the candidates identified. Further research should include independent screening studies, pathway analysis, protein-protein interactions and experimental validation of selected pathway components in order to define the host dependency. In general, this study offers a preliminary analytical basis for the identification and prioritization of potential coronavirus-associated host factors to fuel additional research into their possible significance for host-directed antiviral research.

References

1. Baggen, J., Jacquemyn, M., Persoons, L., Vanstreels, E., Pye, V. E., Wrobel, A. G., ... & Daelemans, D. (2023). TMEM106B is a receptor mediating ACE2-independent SARS-CoV-2 cell entry. Cell, 186(16), 3427-3442.
2. Baggen, J., Persoons, L., Vanstreels, E., Jansen, S., Van Looveren, D., Boeckx, B., ... & Daelemans, D. (2021). Genome-wide CRISPR screening identifies TMEM106B as a proviral host factor for SARS-CoV-2. Nature genetics, 53(4), 435-444.
3. Baggen, J., Vanstreels, E., Jansen, S., & Daelemans, D. (2021). Cellular host factors for SARS-CoV-2 infection. Nature microbiology, 6(10), 1219-1232.
4. Biering, S. B., Sarnik, S. A., Wang, E., Zengel, J. R., Leist, S. R., Schäfer, A., ... & Hsu, P. D. (2022). Genome-wide bidirectional CRISPR screens identify mucins as host factors modulating SARS-CoV-2 infection. Nature Genetics, 54(8), 1078-1089.
5. Brown, K. R., Chan, K., & Moffat, J. (2022). Systematic genome-scale identification of host factors for SARS-CoV-2 infection across models yields a core single gene dependency; ACE2 [Data set]. NCBI Gene Expression Omnibus. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE178699
6. Chan, K., Farias, A. G., Lee, H., Guvenc, F., Mero, P., Aulakh, K., ... & Moffat, J. (2021). Systematic genome-scale identification of host factors for sars-cov-2 infection across models yields a core single gene dependency; ACE2. BioRxiv, 2021-06.
7. Cui, Z., Wang, H., Dong, Y., Liu, S. L., & Wang, Q. (2023). Deciphering and targeting host factors to counteract SARS-CoV-2 and coronavirus infections: insights from CRISPR approaches. Frontiers in genome editing, 5, 1231656.
8. Daniloski, Z., Jordan, T. X., Wessels, H. H., Hoagland, D. A., Kasela, S., Legut, M., ... & Sanjana, N. E. (2021). Identification of required host factors for SARS-CoV-2 infection in human cells. Cell, 184(1), 92-105.

9. Ghosh, S., Dellibovi-Ragheb, T. A., Kerviel, A., Pak, E., Qiu, Q., Fisher, M., ... & Altan-Bonnet, N. (2020). β-Coronaviruses use lysosomes for egress instead of the biosynthetic secretory pathway. Cell, 183(6), 1520-1535.
10. Gordon, D. E., Jang, G. M., Bouhaddou, M., Xu, J., Obernier, K., White, K. M., ... & Krogan, N. J. (2020). A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature, 583(7816), 459-468.
11. Grodzki, M., Bluhm, A. P., Schaefer, M., Tagmount, A., Russo, M., Sobh, A., ... & Norris, M. H. (2022). Genome-scale CRISPR screens identify host factors that promote human coronavirus infection. Genome medicine, 14(1), 10.
12. Hoffmann, H. H., Sánchez-Rivera, F. J., Schneider, W. M., Luna, J. M., Soto-Feliciano, Y. M., Ashbrook, A. W., ... & Rice, C. M. (2021). Functional interrogation of a SARS-CoV-2 host protein interactome identifies unique and shared coronavirus host factors. Cell host & microbe, 29(2), 267-280.
13. Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., ... & Pöhlmann, S. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. cell, 181(2), 271-280.
14. Hou, J., Wei, Y., Zou, J., Jaffery, R., Sun, L., Liang, S., ... & Peng, W. (2024). Integrated multi-omics analyses identify anti-viral host factors and pathways controlling SARS-CoV-2 infection. Nature communications, 15(1), 109.
15. Israeli, M. A., Finkel, Y., Yahalom-Ronen, Y., Paran, N., Chitlaru, T., Israeli, O., ... & Bercovich-Kinori, A. (2022). Genome-wide CRISPR screens identify GATA6 as a proviral host factor for SARS-CoV-2 via modulation of ACE2. Nature communications, 13(1), 2237.
16. Kim, D., Lee, J. Y., Yang, J. S., Kim, J. W., Kim, V. N., & Chang, H. (2020). The architecture of SARS-CoV-2 transcriptome. cell, 181(4), 914-921.
17. Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., ... & Wang, X. (2020). Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. nature, 581(7807), 215-220.
18. Mac Kain, A., Maarifi, G., Aicher, S. M., Arhel, N., Baidaliuk, A., Munier, S., ... & Roesch, F. (2022). Identification of DAXX as a restriction factor of SARS-CoV-2 through a CRISPR/Cas9 screen. Nature communications, 13(1), 2442.
19. May, D. G., Martin-Sancho, L., Anschau, V., Liu, S., Chrisopulos, R. J., Scott, K. L., ... & Roux, K. J. (2022). A BioID-derived proximity interactome for SARS-CoV-2 proteins. Viruses, 14(3), 611.
20. Rebendenne, A., Roy, P., Bonaventure, B., Chaves Valadão, A. L., Desmarets, L., Arnaud-Arnould, M., ... & Goujon, C. (2022). Bidirectional genome-wide CRISPR screens reveal host factors regulating SARS-CoV-2, MERS-CoV and seasonal HCoVs. Nature genetics, 54(8), 1090-1102.
21. Schmidt, N., Lareau, C. A., Keshishian, H., Ganskih, S., Schneider, C., Hennig, T., ... & Munschauer, M. (2021). The SARS-CoV-2 RNA–protein interactome in infected human cells. Nature microbiology, 6(3), 339-353.
22. Schneider, W. M., Luna, J. M., Hoffmann, H. H., Sánchez-Rivera, F. J., Leal, A. A., Ashbrook, A. W., ... & Poirier, J. T. (2021). Genome-scale identification of SARS-CoV-2 and pan-coronavirus host factor networks. Cell, 184(1), 120-132.
23. See, W. R., Yousefi, M., & Ooi, Y. S. (2024). A review of virus host factor discovery using CRISPR screening. MBio, 15(11), e03205-23.
24. Snijder, E. J., Limpens, R. W., de Wilde, A. H., de Jong, A. W., Zevenhoven-Dobbe, J. C., Maier, H. J., ... & Bárcena, M. (2020). A unifying structural and functional model of the coronavirus replication organelle: Tracking down RNA synthesis. PLoS biology, 18(6), e3000715.
25. Ugalde, A. P., Bretones, G., Rodriguez, D., Quesada, V., Llorente, F., Fernández-Delgado, R., ... & López-Otín, C. (2022). Autophagy-linked plasma and lysosomal membrane protein PLAC8 is a key host factor for SARS-CoV-2 entry into human cells. The EMBO journal, 41(21), EMBJ2022110727.
26. V'kovski, P., Kratzel, A., Steiner, S., Stalder, H., & Thiel, V. (2021). Coronavirus biology and replication: implications for SARS-CoV-2. Nature Reviews Microbiology, 19(3), 155-170.
27. Wang, R., Simoneau, C. R., Kulsuptrakul, J., Bouhaddou, M., Travisano, K. A., Hayashi, J. M., ... & Puschnik, A. S. (2021). Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses. Cell, 184(1), 106-119.
28. Wei, J., Alfajaro, M. M., DeWeirdt, P. C., Hanna, R. E., Lu-Culligan, W. J., Cai, W. L., ... & Wilen, C. B. (2021). Genome-wide CRISPR screens reveal host factors critical for SARS-CoV-2 infection. Cell, 184(1), 76-91.
29. Wolff, G., Limpens, R. W., Zevenhoven-Dobbe, J. C., Laugks, U., Zheng, S., de Jong, A. W., ... &

Bárcena, M. (2020). A molecular pore spans the double membrane of the coronavirus replication organelle. Science, 369(6509), 1395-1398.
30. Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., & Zhou, Q. (2020). Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. Science, 367(6485), 1444-1448.
31. Zhu, S., Liu, Y., Zhou, Z., Zhang, Z., Xiao, X., Liu, Z., ... & Wei, W. (2022). Genome-wide CRISPR activation screen identifies candidate receptors for SARS-CoV-2 entry. Science China Life Sciences, 65(4), 701-717.
32. Zhu, Y., Feng, F., Hu, G., Wang, Y., Yu, Y., Zhu, Y., ... & Zhang, R. (2021). A genome-wide CRISPR screen identifies host factors that regulate SARS-CoV-2 entry. Nature communications, 12(1), 961.
33. Ziegler, C. G., Allon, S. J., Nyquist, S. K., Mbano, I. M., Miao, V. N., Tzouanas, C. N., ... & Zhang, K. (2020). SARS-CoV-2 receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues. Cell, 181(5), 1016-1035.